PORTFOLIO VISUALIZATION (PVIZ) DEMO

Krishna Collie and Calvin Johnson

Portfolio Analysis Workshop

February 6, 2012

CIT
Center for Information Technology

DCB

Division of Computational Bioscience
Center for Information Technology | National Institutes of Health

HPCIO

CIT DIVISION OF COMPUTATIONAL BIOSCIENCE

DCB

Division of Computational Bioscience

Center for Information Technology | National Institutes of Health

HPCIO

- The Division of Computational Bioscience (DCB) of CIT is a research and development organization that provides scientific and technical expertise in applying state-of-the-art technologies to support the NIH Intramural Research Program (IRP). Working with NIH's Institutes and Centers (ICs), DCB develops leading-edge computational methods and tools to solve complex biomedical laboratory and clinical research problems.

- The High Performance Computing and Informatics Office responds to critical demands to develop capabilities in new areas, including:

  - Portfolio analysis
  - Text mining and analytics
  - Machine learning
  - Knowledge-based systems
  - High-end computing
  - Scientific visualization

CIT
Center for Information Technology

WHY VISUALIZE THE NIH PORTFOLIO?

- What questions do we want to ask?

  - What do we want to measure?

- Public health impact

  - How has NIH funding affected health?

- Research management

  - Decision support tools, spending trends, ROI

- Discovery of new areas of research

  - novel insights, lead to new programs(RFA/PA)

CIT
Center for Information Technology

PVIZ COMPONENTS

- Visualization of categorical information
  - Tree Map – display hierarchical data
  - Aduna – visualize overlaps between categories
- Exploratory analysis and visualization
  - Circles – visualization of Lingo3G clustering
  - FoamTree – alternative visualization of Lingo3G clustering
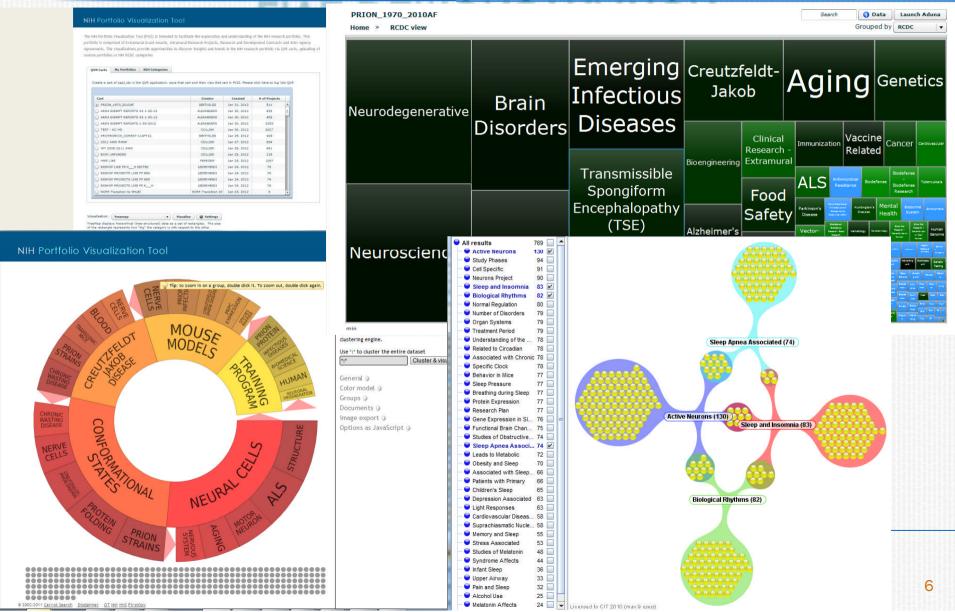 - Aduna – overlaps between clusters

DATA SOURCES AND MANAGEMENT

- ## NIH project types
  - Extramural grant awards
  - Intramural research projects
  - Research and development contracts
  - Inter Agency Agreements
- ## Database
  - IMPACII
- ## FY 2007 – FY2011

- ## Integrated QVR carts
  - Build cart in QVR, visualize in PVIZ
- ## Custom portfolios
  - Upload projects from IC specific data set
- ## RCDC categories
  - Explore enterprise data

CIT
Center for Information Technology

LIVE DEMONSTRATION

PVIZ FUTURE ENHANCEMENTS

- Visualizing similarity of documents and corpora
  - Multiple similarity metrics (e.g. cosine, Jaccard)
  - Heatmap, circle chart, dendrogram

- Incorporating additional clustering algorithms
  - Latent Dirichlet Allocation (LDA), Kohonen maps.

- Classification, validation, and annotation workflows

- Integrating additional data sources
  - , e.g. literature (Thompson Reuters WOS), patents (USPTO)

CIT
Center for Information Technology

THANKS TO: